Qeios

Peer Review

Review of: "A Simple Preprocessing Method Enhances Machine Learning Application to EEG Data for Differential Diagnosis of Autism"

Florencia Maria lucia Sanabria[1]

1. Instituto HELMAR, Argentina

Review of “A Simple Preprocessing Method Enhances Machine Learning Application to EEG Data for Differential Diagnosis of Autism”

The manuscript presents an interesting approach to summarizing EEG recordings through Manhattan distances and minimum spanning trees, followed by machine learning classification. Including children with other neuropsychiatric disorders, rather than only typically developing controls, is a valuable feature.

However, important inconsistencies in the modeling procedure, feature representation, and performance reporting currently limit interpretation. The findings should be presented as preliminary classification results rather than evidence of an established diagnostic biomarker.

Clarify the clinical reference standard and participant selection

The manuscript should explain how autism and comparison diagnoses were established, which diagnostic criteria and assessment procedures were used, and whether diagnostic assessment was independent of EEG findings.

Please report recruitment procedures, medication exposure, intellectual and language abilities, coexisting conditions, and the matching method. Sharing an age range does not itself demonstrate equivalent age distributions.

The comparator group is clinically heterogeneous, including anxiety, sleep disorders, ADHD, mood disorders, oppositional defiant disorder, and traumatic brain injury. The authors should acknowledge that

doi.org/10.32388/PBKC7P

an aggregate classification result cannot establish discrimination against each individual condition.

Excluding epilepsy, genetic conditions, and cerebral malformations also restricts applicability to the broader population undergoing neurodevelopmental assessment.

Reconcile the classifier and validation descriptions

The abstract and results identify KNN as the best classifier, but the validation section describes artificial neural networks, frozen weight matrices, and a second network with identical architecture. This does not adequately describe KNN training or prediction.

The authors should provide the actual procedure used for every reported model, including software, hyperparameters, distance metric, feature scaling, random seeds, and model-selection strategy.

All data-dependent preprocessing, feature selection, and hyperparameter optimization should occur within training partitions. If multiple algorithms were compared, the test partitions should not also serve to select the winning model. Reporting should follow a clinical prediction-model framework such as TRIPOD+AI.

Justify the numerical encoding of electrode identity

Electrode identifiers are anatomical labels, not measurements on a continuous numerical scale. Assigning integers from 1 to 19 and using these values in a distance-based classifier may introduce arbitrary relationships between electrodes.

The authors should explain how electrode identities were encoded and demonstrate that predictions are robust to arbitrary renumbering while anatomical identity is preserved.

Sorting electrodes by degree also requires a reproducible rule for ties. Because many nodes can have the same degree, tie handling could materially affect the feature vector. Comparing the proposed representation with node degrees stored in a fixed anatomical order would be informative.

Describe accurately what the 38 features preserve

A minimum spanning tree with 19 nodes contains 18 edges, and its node degrees sum to 36. The proposed representation records degree information and electrode ordering, but does not retain the complete edge structure or edge weights.

Different trees can therefore produce the same degree-based representation. The statement that all EEG content is transformed into 38 numbers should be replaced with a description of a compressed feature summary.

The manuscript should assess which information is retained, what is discarded, and whether this representation performs better than simpler alternatives.

Strengthen EEG processing and biological interpretation

Essential acquisition and preprocessing details are missing, including sampling rate, frequency filters, amplitude units, artifact-rejection criteria, normalization, and management of drowsiness.

The claim that each recording contains ten continuous artifact-free minutes requires explanation, particularly in children aged four to ten. Please describe who selected the segments and whether the selection was blinded to diagnosis.

Manhattan distance measures differences between recorded signals. Larger distances do not directly establish that the underlying brain regions are more disconnected. Signal amplitude, shared sources, spatial mixing, and vigilance may influence the measurements.

The proposed features may have predictive value without constituting direct measures of neural connectivity; these claims should be separated.

Make the performance estimates auditable.

Provide confusion-matrix counts, denominators, partition-specific results, and confidence intervals. The reported sensitivity of 92.37% and specificity of 94.03% are not readily explained by the stated two-fold procedure involving 50 participants per class. If these values reflect additional repetitions or averaging, that procedure must be reported.

The arithmetic mean of sensitivity and specificity is balanced accuracy. Although it coincides with conventional accuracy in an equally balanced sample under the appropriate aggregation, the terminology and calculation should be explicit.

A single two-fold evaluation provides limited evidence of performance stability. External validation remains necessary, and the validation procedure should not be described as ensuring generalizability.

Clarify TWIST selection and exploratory clustering.

The clustering section refers to features selected by TWIST, but the feature-selection procedure is not described. Explain whether TWIST was used, whether it incorporated diagnostic labels, and whether selection occurred independently within training partitions.

If the embedding uses features selected through diagnostic labels, its apparent separation cannot be presented as wholly independent unsupervised confirmation.

The two-dimensional visualization should remain exploratory and should not substitute for validated classification performance.

Demonstrate the claimed improvement and moderate clinical conclusions.

The title states that preprocessing enhances machine learning, but this requires direct comparisons using identical validation procedures. Include suitable baseline features, alternative representations, and computational benchmarks.

The novelty claim should also be narrowed: minimum spanning trees have previously been applied to brain-network analysis. The contribution may lie in this particular distance-based representation rather than introducing MST analysis to biomedicine.

Finally, the assertion that epilepsy is commonly misdiagnosed as autism requires stronger evidence and should distinguish differential diagnosis from coexistence. Clinical guidance recommends EEG when epilepsy is suspected; the present study does not establish EEG as a replacement for comprehensive autism assessment.

Overall recommendation: Major revision.

This is a potentially useful exploratory study, particularly because it includes clinical comparators. Its contribution depends on a reproducible modeling pipeline, defensible feature encoding, transparent performance calculations, and independent validation. Additional methodological review by an EEG specialist and a statistical or machine learning reviewer would strengthen the work.

Declarations

Potential competing interests: No potential competing interests to declare.